Yule process sample path asymptotics

Arnaud de la Fortelle

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Yule process sample path asymptotics

Arnaud de La Fortelle

N° 5577

May 17, 2005

THÈME 1

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--5577--FR+ENG

Yule process sample path asymptotics

Arnaud de La Fortelle*

Thème 1 — Réseaux et systèmes
Projet PREVAL

Rapport de recherche n° 5577 — May 17, 2005 — 14 pages

Abstract: This research report presents two results on sample paths for the Yule process: one fluid limit theorem and one sample path large deviation result. The main interest does not lie in results by themselves but in the understanding the change of measure gives on the way large deviation occurs in the case of non-homogeneous processes.

Two different phenomena, with respect to classical large deviations principles, are exhibited. First, the probability decay rate has the same form whatever the speed of divergence from the standard behavior. Second, a large deviation event does not take place by "twisting" constantly the transition rates but this deformation is concentrated on an infinitely small proportion of the transition, yet on an infinite number of transitions!

Key-words: Large deviations, random trees, branching process, fluid limit, Yule process, martingale, change of measure.

* Postal address: INRIA - Domaine de Voluceau, Rocquencourt - BP 105 - 78153 Le Chesnay Cedex - FRANCE. e-mail address: `Arnaud.De_La_Fortelle@inria.fr`

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 Le Chesnay Cedex (France)
Téléphone : +33 1 39 63 55 11 — Télécopie : +33 1 39 63 53 30

Asymptotiques trajectorielles pour le processus de Yule

Résumé : Ce rapport de recherche présente deux résultats trajectoriels sur les processus de Yule : un théorème de limite fluide et un résultat de grandes déviations. Le principal intérêt ne réside pas dans les théorèmes eux-mêmes mais dans la compréhension que le changement de mesure donne de la manière dont les grandes déviations se construisent dans ce cas simple de processus non-homogène.

On montre deux phénomèmes différents des résultats classiques de grandes déviations. Le premier est que la forme de la décroissance des probabilités est la même quelle que soit la vitesse à laquelle on s'écarte du comportement "moyen". Le deuxième est qu'une grande déviation ne se produit pas en "tordant" de façon constante les intensités d'arrivée, mais en concentrant cette déformation sur une proportion infiniment faible de transitions (mais un nombre infini !).

Mots-clés : Grandes déviations, arbres aléatoires, processus de branchement, limite fluide, processus de Yule, martingale, changement de mesure

1 Introduction

1.1 The model

This paper deals with asymptotics for the Yule process which is defined as follows: $\{Y_t(n), t \geq 0\}$ is a pure birth Markov process with initial state $Y_0(n) = n$ and birth rate $\lambda Y_t(n)$ at time t. When n is not specified, $Y_t \stackrel{\text{def}}{=} Y_t(1)$.

The Yule process is a model for many phenomenon and pertains to several queuing systems and branching processes. For example, Y_t represents the total number of nodes for random trees without deletions: the tree begins with a single root and then new nodes arrive at all nodes with intensity λ. For a classical introduction to branching processes, see [1].

The main motivation of this paper was to understand the asymptotic behavior of random trees as in [8]. However, asymptotics of non-homogeneous processes is quite difficult and we want to understand first how simple processes behave. The fluid limits and large deviations asymptotics presented here show new phenomenon that we believe are shared by many processes similar to the Yule process, including random trees. For a presentation of large deviations we refer to [7] and references therein.

A great amount of research has been done around the Yule process, mostly related to Binary Search Trees (BST). For example, the martingale defined in (3.2) can be linked to the so-called BST martingale (see [15]). Some references are provided at the end of the paper. However, we analyze specific features which, up to our knowledge, are not discussed in the literature.

Interesting peculiarities of the large deviations of the Yule process presented here, that we think are common with other branching processes, is first that there is no "typical" speed of large deviation, and second that the large deviation event is not built on the whole interval $[0, T]$ for large T, but the process is "twisted" on a finite interval of time at the beginning, and then continues quite "normally".

1.2 Yule process and related processes

The Yule process is closely connected to another non-homogeneous process. Let X_n be the sum of n exponential variables $E_1, \ldots, E_n$ with respective parameters $\lambda, \ldots, n\lambda$:

$$X_n \stackrel{\text{def}}{=} E_1 + \cdots + E_n. \tag{1.1}$$

The relationship between Y_t and X_n is given by

$$X_n \geq t \iff Y_t \leq n \qquad \forall t \geq 0, \forall n \geq 1. \tag{1.2}$$

We will see that considering X_n is helpful when heuristically calculating changes of measures for Y_t.

The distribution of Y_t is a geometric one

$$\mathbb{P}\left[Y_t = k\right] = e^{-\lambda t}\left(1 - e^{-\lambda t}\right)^{k-1}, \qquad \text{if } k \geq 1. \tag{1.3}$$

The distribution of $Y_t(n)$ is equal in law to the sum of n independent Yule process $Y_t(1)$ hence the distribution is a negative binomial

$$\mathbb{P}\left[Y_t(n) = k\right] = \binom{n-1}{k-1} e^{-n\lambda t}\left(1 - e^{-\lambda t}\right)^{k-1}, \qquad \text{if } k \geq n. \tag{1.4}$$

Hence there is no major challenge in the calculation of the large deviations decay rate for the Yule process. Indeed, the probability to grow faster than "usual" is easily calculated for large t, e.g.

$$\mathbb{P}\left[Y_t \geq e^{a\lambda t}\right] \simeq \exp\{-e^{(a-1)\lambda t}\} \qquad \forall a > 0, \tag{1.5}$$

hence the decay rate is $\exp\{(a-1)\lambda t\}$, far greater (for $a > 1$) than the usual linear one. Now, this calculation does not give any insight on the way this deviation occurs. This is what we intend to show in this paper.

1.3 Large deviations behavior

The main result of the paper, Proposition 3.1, states roughly that the probability for the Yule process to follow a path $t \to C(e^{\lambda t} - 1)$ for $t \in [0, T]$ is of the order e^{-C}. This is coherent with (1.5): take $C = \exp\{(a-1)\lambda t\}$. Therefore, *there is no typical speed*: the only condition is for $C = C(T)$ to go to infinity with T.

The change of measure produced by the martingale defined in (3.2) shows the typical behavior of such a large deviation event. The transition rate accelerates from $n\lambda$ in state n to $n\lambda + C$. Recall that n evolves from 1 to $Ce^{\lambda T}$. For small n ($n \ll C$), this is a very large deformation and the "cost" (i.e. the contribution to the decay rate C) to do so is very high (the order is $\log C$). For large n ($n \gg C$), the cost is almost null (the order is $(Cn^{-1})^2/2$). These approximations can be derived from equation (A.8).

Therefore, there is only a small fraction of transitions that are significantly changed: taking into account the C first ones, this make a proportion $e^{-\lambda T}$ that is asymptotically null. Turning this proportion of number into proportion of time, one calculates easily that the fluid limit crosses the level C at time $T' = \lambda^{-1}\log 2$ (the solution of $C(e^{\lambda t}-1) = C$). Hence there is a fixed time (T' does not depend on C nor on T) for the large deviations behavior, then the process "continues" with an almost unmodified behavior.

This behavior is understandable if we consider the Yule process as the number of nodes in a tree. It is easier to accelerate the replication of a small number of nodes than a large number. Hence the global acceleration is concentrated on the beginning, when there are few nodes. This explanation is meaningless mathematically (there is no "global acceleration") but this is the idea.

We analyzed only the acceleration case. When coming to the deceleration ($C \to 0$), the behavior is even simpler: the first transition is stopped as long as necessary, then the process goes on almost normally. This case is only sketched in Appendix A as a comment of equation (A.5).

The structure of the report is the following. Appendix A shows how the result can be heuristically found while proofs are given, in Section 2 for the fluid limit and in Section 3 for the large deviations asymptotics.

2 Fluid limit

The Yule process can be considered as the time-reversed of the M/M/∞ queue with no entrance. Therefore many tools are similar. First, it is easily checked that, for all $c > -1$

$$h(x,t) = (1 + ce^{-\lambda t})^{-x}$$

is space-time harmonic w.r.t. the Yule process generator. This means

$$(1 + ce^{-\lambda t})^{-Y_t}$$

is a martingale. The decomposition

$$(1+w)^{-x} = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{(x+n-1)!}{(x-1)!} w^n$$

yields the martingales $e^{-n\lambda t}Y_t(Y_t+1)\dots(Y_t+n-1)$. Most useful are the martingales for $n=1$ and $n=2$:

$$Y_t e^{-\lambda t}, \tag{2.1}$$

$$Y_t(Y_t+1)e^{-2\lambda t}. \tag{2.2}$$

It is easily checked that these martingales are uniformly bounded, hence converge a.s. and in L_1. It is well known that the first one converges to an exponential r.v. with parameter 1 if $Y_0=1$. Since the law of $Y_t(n)$ is equal to the sum of n independent Yule processes beginning at 0, the martingale $Y_t(n)e^{-\lambda t}$ converges to the sum of n i.i.d. exponential r.v. with parameter 1.

The fluid limit is obtained by applying Doob's inequality to the martingale $Y_t(n)e^{-\lambda t}-n$.

$$\mathbb{P}\left[\sup_{t\le T}|Y_t(n)e^{-\lambda t}-n|\ge n^\alpha\right]\le\frac{\mathbb{E}\left[|Y_T(n)e^{-\lambda T}-n|\right]}{n^\alpha}.$$

Using $\mathbb{E}\left[|X|\right]\le\sqrt{\mathbb{E}\left[X^2\right]}$ and the martingales (2.1) and (2.2), one gets

$$\mathbb{P}\left[\sup_{t\le T}|Y_t(n)e^{-\lambda t}-n|\ge n^\alpha\right]\le n^{\frac12-\alpha}. \tag{2.3}$$

This inequality is stronger than the usual fluid limit formulation, where one classically chooses $\alpha=1$. Of course this does not look like classical fluid limits, since the scaling is not Euler's scaling, but this is precisely what is necessary for the large deviations asymptotics.

Moreover, one can show that the scaled process $Z_t\stackrel{\text{def}}{=}n^{-1}(Y_t(n)-n)$ converges to the fluid equation $y'=\lambda(1+y)$ with initial condition $y(0)=0$. The solution is $y(t)=e^{\lambda t}-1$. On a fluid limit scale, we should then have $Y_t(n)\simeq n(e^{\lambda t}-1)+n=ne^{\lambda t}$. This is precisely what states Proposition 2.1.

Proposition 2.1 (Fluid limit) *For all $\alpha>1/2$, for all $T\ge0$,*

$$\lim_{n\to\infty}\mathbb{P}\left[\sup_{t\le T}|Y_t(n)e^{-\lambda t}-n|\ge n^\alpha\right]=0.$$

Define $X_k(n)\stackrel{\text{def}}{=}E_n+\cdots+E_{n+k}$, which is the "image" of $Y_t(n)$. Then Proposition 2.1 can be translated to $X_k(n)$ via the relation (1.2). More precisely, (2.3) becomes

$$\mathbb{P}\left[\sup_{n\le k\le e^{\lambda T}}\left|X_k(n)-\lambda^{-1}\log\left(\frac{k}{n}\right)\right|\ge\lambda^{-1}n^{\alpha-1}\right]\le n^{\frac12-\alpha}. \tag{2.4}$$

3 Large deviations

3.1 Change of Measure

Following the heuristics of Appendix A, for a given constant $C \geq 0$, we set

$$f(n) \stackrel{\text{def}}{=} \sum_{i=1}^{n} \log \frac{i+C}{i}, \qquad \forall n \geq 0. \tag{3.1}$$

and we define the process

$$M_t \stackrel{\text{def}}{=} \exp\left\{f(Y_t - 1) - \lambda C t\right\}. \tag{3.2}$$

One checks that M_t is a martingale

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{E}\left[M_{t+s}|\mathcal{F}_t\right] &= \left(\lambda Y_t \left(e^{f(Y_t)-f(Y_t-1)} - 1\right) - \lambda C\right) \exp\{f(Y_t-1) - \lambda C t\} \\
&= \left(\lambda Y_t \left(\frac{C+Y_t}{Y_t} - 1\right) - \lambda C\right) M_t = 0
\end{aligned}$$

It is then classical to define a new probability measure, the "twisted" probability, by

$$\mathbb{P}^*\left[A\right] \stackrel{\text{def}}{=} \mathbb{E}\left[\mathbb{1}_{\{A\}} M_t\right], \qquad \forall A \in \mathcal{F}_t.$$

Under this twisted probability, the process Y_t is still Markovian with intensity $\lambda(C + Y_t)$: for any bounded function h,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{E}^*\left[h(Y_{t+s})|\mathcal{F}_t\right] &= M_t^{-1}\frac{\mathrm{d}}{\mathrm{d}s}\mathbb{E}\left[h(Y_{t+s}M_{t+s})|\mathcal{F}_t\right] \\
&= \lambda Y_t \left(e^{f(Y_t+1)-f(Y_t)} h(Y_t+1) - h(Y_t)\right) - \lambda C h(Y_t) \\
&= \lambda(C+Y_t)\left(h(Y_t+1) - h(Y_t)\right))
\end{aligned}$$

This means that, under $\mathbb{P}^*$, $Y_t(1)$ has the same law as $Y_t(C+1) - C$ under $\mathbb{P}$. As will be useful later, this yields, combining with the fluid inequality 2.3:

$$\begin{aligned}
&\mathbb{P}^*\left[\sup_{t\leq T} |(Y_t(1)+C)e^{-\lambda t} - C - 1| < (C+1)^\alpha\right] \qquad (3.3) \\
&= \mathbb{P}\left[\sup_{t\leq T} |Y_t(C+1)e^{-\lambda t} - (C+1)| < (C+1)^\alpha\right] \geq 1 - (C+1)^{\frac{1}{2}-\alpha}.
\end{aligned}$$

For the sake of simplicity, we denote the large deviation event by

$$E_{\mathrm{ld}}(T, C, \alpha) \stackrel{\text{def}}{=} \left\{\sup_{t\leq T} |(Y_t(1)+C)e^{-\lambda t} - C - 1| < (C+1)^\alpha\right\} \tag{3.4}$$

3.2 Lower bound

Since $M_t > 0$ a.s, $\mathbb{P}$ and $\mathbb{P}^*$ are reciprocally absolutely continuous and $\mathbb{P}[A] = \mathbb{E}^*\left[\mathbb{1}_{\{A\}} M_t^{-1}\right]$ for all $A \in \mathcal{F}_t$. Therefore we derive from , for any integer $C \geq 0$ and any real $T > 0$

$$\mathbb{P}\left[E_{\mathrm{ld}}(T,C,\alpha)\right] = \mathbb{E}^*\left[\mathbb{1}_{\{E_{\mathrm{ld}}(T,C,\alpha)\}} M_T^{-1}\right] \geq \mathbb{P}^*\left[E_{\mathrm{ld}}(T,C,\alpha)\right] \inf_{E_{\mathrm{ld}}(T,C,\alpha)} M_T^{-1} \quad (3.5)$$

The first term is bounded by (3.3). Bounds on the second term, the infimum, come from the bounding of f: inequalities A.16. Indeed, on the event $E_{\mathrm{ld}}(T,C,\alpha)$,

$$(C+1-(C+1)^\alpha)e^{\lambda T} - C < Y_T < (C+1+(C+1)^\alpha)e^{\lambda T} - C \quad (3.6)$$

Combining with inequalities A.16 yields

$$f(Y_T - 1) \leq C \log \frac{(C+1+(C+1)^\alpha)e^{\lambda T}}{C} + C \leq \lambda C T + 1 + (C+1)^\alpha + C. \quad (3.7)$$

This immediately gives the following bound on M_T

$$\inf_{E_{\mathrm{ld}}(T,C,\alpha)} M_T^{-1} \geq \exp\left\{-(C+1) - (C+1)^\alpha\right\} \quad (3.8)$$

Finally we get the large deviations lower bound

$$\log \mathbb{P}\left[E_{\mathrm{ld}}(T,C,\alpha)\right] \geq -(C+1) - (C+1)^\alpha + \log\bigl(1 - (C+1)^{\frac{1}{2}-\alpha}\bigr). \quad (3.9)$$

3.3 Upper bound

The large deviations upper bound is obtained by reversing the equation (3.5). Since the probability is bounded by 1:

$$\mathbb{P}\left[E_{\mathrm{ld}}(T,C,\alpha)\right] \geq \sup_{E_{\mathrm{ld}}(T,C,\alpha)} M_T^{-1} \quad (3.10)$$

Combining, as for the lower bound, (3.6) with inequalities (A.16) yields

$$f(Y_T - 1) \geq \lambda C T + C - R_1 \quad (3.11)$$

with the rest (obtained using $\log(1-x) \geq -x/(1-x)$):

$$\begin{aligned} R_1(T,C,\alpha) &\stackrel{\text{def}}{=} C\frac{(C+1)^{\alpha-1}+(C+1)^{-1}e^{-\lambda T}}{1-(C+1)^{\alpha-1}-(C+1)^{-1}e^{-\lambda T}} \\ &\quad + \frac{C^2}{2((C+1-(C+1)^{\alpha})e^{\lambda T}-C-1)} + \log(C+1) \\ &\leq \frac{C(C+1)^{\alpha}}{C+(C+1)^{\alpha}} + \frac{Ce^{-\lambda T}}{2}\frac{1}{1-C^{\alpha-1}-e^{-\lambda T}} + \log(C+1) \\ &\leq C^{\alpha} + Ce^{-\lambda T} + \log(C+1) \end{aligned} \tag{3.12}$$

assuming $C^{\alpha-1} \leq 1/4$ and $e^{-\lambda T} \leq 1/4$ which will be the case since C and T will be taken indefinitely large. Under this condition, we get the large deviations upper bound

$$\log \mathbb{P}\left[E_{\mathrm{ld}}(T,C,\alpha)\right] \leq -C + C^{\alpha} + Ce^{-\lambda T} + \log(C+1). \tag{3.13}$$

3.4 Large deviations

Gathering bounds (3.9) and (3.13), one gets the large deviations asymptotics. More precisely, for all $C : \mathbb{R}_+ \mapsto \mathbb{Z}_+$ tending to infinity, for all $1/2 < \alpha < 1$,

$$\lim_{T\to\infty} \frac{1}{C(T)} \log \mathbb{P}\left[E_{\mathrm{ld}}(T,C(T),\alpha)\right] = -1. \tag{3.14}$$

One can simplify the event $E_{\mathrm{ld}}(T,C,\alpha)$ and keep the same limit. The modified lower bound is obtained by changing α and the upper bound by directly modifying the bounds (3.10)–(3.13). Moreover, we can easily relax the condition for C to be integer. Therefore we get the following proposition.

Proposition 3.1 (Large deviations) *For all $C : \mathbb{R}_+ \mapsto \mathbb{R}_+$ tending to infinity, for all $1/2 < \alpha < 1$,*

$$\lim_{T\to\infty} \frac{1}{C(T)} \log \mathbb{P}\left[\sup_{t\leq T} |(Y_t(1)+C(T))e^{-\lambda t} - C(T)| < C^{\alpha}(T)\right] = -1. \tag{3.15}$$

A sequel of this proposition is, roughly stated, that the probability to reach $Y_T \simeq C(T)e^{\lambda T}$ decreases as $e^{-C(T)}$. Taking $C(T) = e^{(\alpha-1)\lambda T}$, one gets the result (1.5). What we gained with this demonstration is first an explanation of the form of this relationship between increase rate and exponential decrease rate for *any*

$C(T)$ tending to infinity, and second we know the way this large deviation event occurs by mean of the change of measure. In the present case, we could even state the convergence of the conditional probabilities, but the proof would lengthen the paper and does not bring further substantial understanding of the behavior of the process.

Appendix A Heuristics around the decay rate

This section shows some calculations that help to discover the right functions that are then used in the demonstrations. Therefore the emphasis is not on proving, but giving some insight on where these functions come from.

A.1 Optimization

We consider the process X_n. Let define Λ_i the cumulant generating function for E_i and the martingale

$$M_n = \exp\left\{\sum_{i=1}^{n} \alpha_i E_i - \Lambda_i(\alpha_i)\right\} \qquad \text{with } \alpha_i > -\lambda i \quad \forall i. \tag{A.1}$$

The large deviation rate function is obtained by the Legendre transform Λ_i^* of Λ_i, which yields a rate function of the form

$$I \stackrel{\text{def}}{=} \inf_{(t_i)} \sum_{i=1}^{n} \Lambda_i^*(t_i) = \inf_{(t_i)} \sum_{i=1}^{n} -\log(i\lambda t_i) - 1 + i\lambda t_i \tag{A.2}$$

under the constraint

$$\sum_{i=1}^{n} t_i = T. \tag{A.3}$$

Using Lagrange's multipliers, one finds that there exists L such that

$$t_i = \frac{1}{L + i\lambda} \tag{A.4}$$

The transition rates in the twisted process (i.e. such that $X_n = T$ is typical for large n and $T = T(n)$) are

$$\tilde{\lambda}_i = t_i^{-1} = L + \lambda i. \tag{A.5}$$

The values $-1 < L < 0$ correspond to an acceleration of the process X_n, i.e. to a deceleration of Y_t. Heuristically, one sees that this effect is achieved mainly by acting on the first term $1/(1+L)$: one can indeed proves that a deceleration of the Yule process is achieved by stopping the first transition.

The results are very different for $L > 0$, corresponding to an acceleration of Y_t. The first transitions are proportionally more "twisted", but infinitely many as we will see.

In the sequel we will study the case $L > 0$. From equations (A.3) and (A.4), we calculate roughly that $T \simeq \lambda^{-1} \log(\lambda n/L)$ and more precisely

$$\lambda^{-1} \log \frac{\lambda n + L + \lambda}{L + \lambda} \leq T \leq \lambda^{-1} \log \frac{\lambda n + L}{L} \tag{A.6}$$

Let accelerate Y_t, say to $Ce^{\lambda t}$ (instead of $e^{\lambda t}$), with $C = C(t) \to \infty$. This relation on Y_T is transposed to $X_n \simeq \alpha^{-1}\lambda^{-1} \log n = T$ and, combining with the relation $T \simeq \lambda^{-1} \log(\lambda n/L)$, we find

$$L = \lambda C. \tag{A.7}$$

For example, if we want, as in (1.5), $Y_T \simeq e^{a\lambda T}$ with $a > 1$, then $C = e^{(a-1)\lambda T}$. We see the transitions rates are deeply modified for i less or of the same order as C and very slightly modified for $i \gg C$. Therefore there are an infinite number of changed transitions at the beginning, that are very modified, but this is a null proportion of the transitions ($C \ll e^{a\lambda T}$) and the other ones are almost untouched. This is clearly an uncommon large deviations behavior.

For the decay rate, calculations yield (see (A.2) for the definition of Λ_i^*)

$$\Lambda_i^*(t_i) = \log\left(1 + \frac{L}{\lambda i}\right) - \frac{L}{L + \lambda i} = \log\left(1 + \frac{C}{i}\right) - \frac{C}{C + i} \tag{A.8}$$

Using the primitive

$$\int \log\left(1 + \frac{C}{x}\right) - \frac{C}{C + x}\, \mathrm{d}x = x \log\left(1 + \frac{C}{x}\right),$$

and the monotonicity of the integrand, one gets

$$\left[x \log\left(1 + \frac{C}{x}\right)\right]_0^n \leq I \leq \left[x \log\left(1 + \frac{C}{x}\right)\right]_1^{n+1}$$

hence

$$C - \frac{C^2}{2n} \leq I \leq C \tag{A.9}$$

Therefore we get

$$I \simeq C. \tag{A.10}$$

This result is exactly, but heuristically, what is stated in the large deviations Proposition 3.1.

A.2 Martingale

After this optimization step that gives already a clear vision of the twisted behavior of the Yule process, we want to exhibit the martingale that will be used for a rigorous proof.

In the martingale M_n defined in (A.1), the α_i are obtained for a given t_i (replacing formally E_i) by optimization (the Legendre transform):

$$\alpha_i = \lambda i - \frac{1}{t_i} = -L = -\lambda C. \tag{A.11}$$

The good surprise is that α_i does not depend on i. Replacing α_i in (A.1) yields

$$M_n = \exp\left\{-\lambda C X_n - \sum_{i=1}^{n} \log \frac{i}{i+C}\right\} = \exp\left\{f(n) - \lambda C X_n\right\}, \tag{A.12}$$

where f is defined by

$$f(n) \stackrel{\text{def}}{=} \sum_{i=1}^{n} \log \frac{i+C}{i}, \qquad \forall n \geq 0. \tag{A.13}$$

We can then guess the corresponding martingale for Y_t. We change X_n into t and n into $Y_t - 1$ (because formally $X_0 = 0$), so that

$$M_t \stackrel{\text{def}}{=} \exp\left\{f(Y_t - 1) - \lambda C t\right\}. \tag{A.14}$$

One then checks it is indeed a martingale.

The next step is to get bounds for $f(n)$. First, using the monotonicity of the logarithm,

$$\left[x \log x - x\right]_{a-1}^{b} \leq \sum_{i=a}^{b} \log i \leq \left[x \log x - x\right]_{a}^{b+1}, \qquad \forall a, b \geq 1 \tag{A.15}$$

If C is a positive integer,

$$f(n) = \sum_{i=n+1}^{n+C} \log i - \sum_{i=1}^{C} \log i.$$

Therefore, using (A.15) and the inequalities $x - x^2/2 \leq \log(1+x) \leq x$ for $x \geq 0$,

$$C \log \frac{n+C}{C+1} + C - \frac{C^2}{2n} - \log(C+1) \leq f(n) \leq C \log \frac{n+C+1}{C} + C \tag{A.16}$$

So that $f(n) \simeq C \log(n/C) + C$.

References

[1] ATHREYA, K. B., AND NEY, P. E. *Branching processes*. Springer-Verlag, New York, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[2] BIGGINS, D. D. Martingale convergence and large deviations in the branching random walk. *Teor. Veroyatnost. i Primenen. 37*, 2 (1992), 301–306.

[3] BIGGINS, J. D. Large deviations for mixtures. *Electron. Comm. Probab. 9* (2004), 60–71 (electronic).

[4] BIGGINS, J. D., AND BINGHAM, N. H. Large deviations in the supercritical branching process. *Adv. in Appl. Probab. 25*, 4 (1993), 757–772.

[5] BIGGINS, J. D., AND GREY, D. R. A note on the growth of random trees. *Statist. Probab. Lett. 32*, 4 (1997), 339–342.

[6] CHAUVIN, B., DRMOTA, M., AND JABBOUR-HATTAB, J. The profile of binary search trees. *Ann. Appl. Probab. 11*, 4 (2001), 1042–1062.

[7] DEMBO, A., AND ZEITOUNI, O. *Large deviations techniques and applications*, second ed. Springer-Verlag, New York, 1998.

[8] FAYOLLE, G., KRIKUN, M., AND LASGOUTTES, J.-M. Birth and death processes on certain random trees: classification and stationary laws. *Probab. Theory Related Fields 128*, 3 (2004), 386–418.

[9] JABBOUR-HATTAB, J. Martingales and large deviations for binary search trees. *Random Structures Algorithms 19*, 2 (2001), 112–127.

[10] LIU, Q., RIO, E., AND ROUAULT, A. Limit theorems for multiplicative processes. *J. Theoret. Probab. 16*, 4 (2003), 971–1014 (2004).

[11] MAHMOUD, H. M. *Evolution of random search trees*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons Inc., New York, 1992. A Wiley-Interscience Publication.

[12] PITTEL, B. On growing random binary trees. *J. Math. Anal. Appl. 103*, 2 (1984), 461–480.

[13] ROUAULT, A. Large deviations and branching processes. In *Proceedings of the 9th International Summer School on Probability Theory and Mathematical Statistics (Sozopol, 1997)* (2000), vol. 13, pp. 15–38.

[14] Rouault, A. Large deviations for cascades and cascades of large deviations. In *Mathematics and computer science. III*, Trends Math. Birkhäuser, Basel, 2004, pp. 351–362.

[15] Rouault, A., and Chauvin. Connecting yule process, bisection and binary search tree via martingales. *J. of the Iranian Stat. Society 3*, 2 (2004).

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399